Statistical Reflections
Half-day Workshop on Statistical Methods

Venice May 22, 2024

Ca' Foscari University
Scientific Campus
Aula Delta 1B
Via Torino 155
Venezia Mestre

Program:
- 14:00 **David Firth** (University of Warwick): Compositional quasi-likelihood and logit models
- 14:45 **Ioannis Kosmidis** (University of Warwick): Consistent, tuning-free model selection in regression problems
- 15:30 **Omiros Papaspiliopoulos** (Bocconi University): Functional estimation of the marginal likelihood

Professor **David Firth**
University of Warwick, UK
https://warwick.ac.uk/dfirth

Compositional quasi-likelihood and logit models

A composition vector describes the relative sizes of parts of a thing. Some important modern application areas are microbiome analysis, time-use analysis and archaeometry (to name just three). We develop model-based analysis of composition, through the first two moments of measurements on their original scale. In current applied work the most-used route to compositional data analysis, following an approach introduced by the late John Aitchison in the 1980s, is based on contrasts among log-transformed measurements. The quasi-likelihood model framework developed here provides a general alternative with several advantages. These include robustness to secondary aspects of model specification, stability when there are zero-valued or near-zero measurements in the data, and more direct interpretation. Linear models for log-contrast transformed data are replaced by generalized linear models with logit link, and variance-covariance estimation is straightforward via suitably standardized residuals. Joint work with Fiona Sammut, University of Malta.

Professor **Ioannis Kosmidis**
University of Warwick, UK
https://www.ikosmidis.com

Consistent, tuning-free model selection in regression problems

We present a family of consistent model selection procedures for regression problems, which appropriately threshold readily available statistics after estimating the model with all covariate information. The model selection procedures rely only on standard assumptions about information accumulation that guarantee the typically expected asymptotic properties of the estimators (e.g., consistency) and require neither the selection of tuning parameters nor the fitting of all possible models. We apply our thresholding methodology in widely used statistical models (e.g. generalized linear models and beta regression), demonstrate that it is easily implementable as part of the standard maximum likelihood output, and compare its performance to existing techniques, illustrating its effectiveness. Joint work with Patrick Zietkiewicz, University of Warwick.

Professor **Omiros Papaspiliopoulos**
Bocconi University, Italy
https://dec.unibocconi.eu/people/omiros-papaspiliopoulos

Functional estimation of the marginal likelihood

We consider the problem of exploring the marginal likelihood of low-dimensional hyperparameters from high-dimensional Bayesian hierarchical models. We provide a unified framework that connects seemingly unconnected approaches to estimating normalizing constants, including the Vardi estimator, the umbrella sampling and the Gibbs sampler. The framework requires Monte Carlo sampling from the posteriors of the high-dimensional parameters for different values of the hyperparameters on a lattice. We establish a surprising reproducing property that leads to a functional estimator of the marginal likelihood and establish consistency both fixed-lattice and lattice-infill consistency. The resultant method is highly practical, black-box with theoretical guarantees.